

JW



PCT

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/831,683B

DATE: 04/03/2003

TIME: 14:25:52

Input Set : A:\BB1270 USPCT Seq List Rev Feb 03.txt

Output Set: N:\CRF4\04032003\I831683B.raw

3 <110> APPLICANT: E. I. du Pont de Nemours and Company
 5 <120> TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetases
 7 <130> FILE REFERENCE: BB1270
 9 <140> CURRENT APPLICATION NUMBER: US/09/831,683B
 C--> 10 <141> **CURRENT FILING DATE: 2003-02-20**
 12 <150> PRIOR APPLICATION NUMBER: 60/107,789
 13 <151> PRIOR FILING DATE: 1998-11-10
 15 <160> NUMBER OF SEQ ID NOS: 38
 17 <170> SOFTWARE: Microsoft Office 97
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 1178
 21 <212> TYPE: DNA
 22 <213> ORGANISM: Zea mays
 24 <400> SEQUENCE: 1

25	gcacgaggtt	tctataatcc	ttatatctcct	caagtgcctgg	aggaattgag	taacaaaggc	60									
26	ttgatcaagg	agagtgaggg	tgcccgagtt	atattttattc	aagggtcatca	aatccctttg	120									
27	attgttggtta	agagtgatgg	tggcttcaac	tatgcctcaa	cagacttaac	tgctctttgg	180									
28	tatcggtcga	atgttgagca	ggcagagtgg	atcatatatg	ttacagatgt	tggtcagcag	240									
29	cagcactttg	acatggtttt	cagtgcctgca	aagatggccg	gttggctccc	agatccaagt	300									
30	gaaaagaagt	ttccgaaaac	aagccatggt	ggatttggtc	ttgttcttgg	ttcagatggc	360									
31	aagcgggttc	gaaccgcgag	tactgaggtt	gttcgattgg	tagagctact	tgatgaggct	420									
32	aatctcggga	gcaaatcaga	actactacaa	cggctcactg	aaaatggcaa	aattggtgac	480									
33	tggacgggatg	aggaattaga	gcaaacttca	gaggctgttg	gatatggtgc	tgtgaagtac	540									
34	gctgatctaa	aaaataacag	gctcactaat	tacacattta	gttttgaaca	aatgctgagc	600									
35	gataagggaa	atactgctgt	gtaccttcag	tatgcacatg	ctcgtatttg	ttccattatt	660									
36	cggaaatcca	acaagaacgt	ggaagagctg	aagatgagtg	gagccatttc	tctcgaccat	720									
37	ccggatgagc	gcgtgttggg	gotgtatctt	atccgatttg	cagagggttg	tgaagaggca	780									
38	tgcacgaatc	tacttccaaa	tggtgtgtgt	gaatacttgt	acaatctatc	tgaaatgttc	840									
39	acaaaattct	ataccaactg	ccagggtggt	gggtcgccgg	aggagacgag	ccggttggtg	900									
40	ctttgccagg	cgactgctgt	tgtcatgcca	cagtgcctca	acctgctcgg	gatcacgcca	960									
41	gtatacaagc	tgtgattggc	tgcattgttcg	attaatacat	tcaacatgta	gaaaccccaa	1020									
42	ttcatcatgg	ttgcagtttt	ggtcttgtaa	cctagttgag	gcagttaaca	taatctactg	1080									
43	tcctgtttga	aaacagaagg	aactcaaaag	gttgatataa	aatgtgcttg	cagagtttct	1140									
44	gttactaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaa			1178									
46	<210>	SEQ ID NO: 2														
47	<211>	LENGTH: 321														
48	<212>	TYPE: PRT														
49	<213>	ORGANISM: Zea mays														
51	<400>	SEQUENCE: 2														
52	Phe	Tyr	Asn	Pro	Tyr	Ile	Pro	Gln	Val	Leu	Glu	Glu	Leu	Ser	Asn	Lys
53	1					5									10	15
55	Gly	Leu	Ile	Lys	Glu	Ser	Glu	Gly	Ala	Arg	Val	Ile	Phe	Ile	Gln	Gly
56						20				25					30	

P.6
ENTERED

RAW SEQUENCE LISTING

DATE: 04/03/2003

PATENT APPLICATION: US/09/831,683B

TIME: 14:25:52

Input Set : A:\BB1270 USPCT Seq List Rev Feb 03.txt

Output Set: N:\CRF4\04032003\I831683B.raw

```

58 His Gln Ile Pro Leu Ile Val Val Lys Ser Asp Gly Gly Phe Asn Tyr
59          35          40          45
61 Ala Ser Thr Asp Leu Thr Ala Leu Trp Tyr Arg Leu Asn Val Glu Gln
62          50          55          60
64 Ala Glu Trp Ile Ile Tyr Val Thr Asp Val Gly Gln Gln Gln His Phe
65 65          70          75          80
67 Asp Met Val Phe Ser Ala Ala Lys Met Ala Gly Trp Leu Pro Asp Pro
68          85          90          95
70 Ser Glu Lys Lys Phe Pro Lys Thr Ser His Val Gly Phe Gly Leu Val
71          100          105          110
73 Leu Gly Ser Asp Gly Lys Arg Phe Arg Thr Arg Ser Thr Glu Val Val
74          115          120          125
76 Arg Leu Val Glu Leu Leu Asp Glu Ala Lys Ser Arg Ser Lys Ser Glu
77          130          135          140
79 Leu Leu Gln Arg Leu Thr Glu Asn Gly Lys Ile Val Asp Trp Thr Asp
80 145          150          155          160
82 Glu Glu Leu Glu Gln Thr Ser Glu Ala Val Gly Tyr Gly Ala Val Lys
83          165          170          175
85 Tyr Ala Asp Leu Lys Asn Asn Arg Leu Thr Asn Tyr Thr Phe Ser Phe
86          180          185          190
88 Glu Gln Met Leu Ser Asp Lys Gly Asn Thr Ala Val Tyr Leu Gln Tyr
89          195          200          205
91 Ala His Ala Arg Ile Cys Ser Ile Ile Arg Lys Ser Asn Lys Asn Val
92          210          215          220
94 Glu Glu Leu Lys Met Ser Gly Ala Ile Ser Leu Asp His Pro Asp Glu
95 225          230          235          240
97 Arg Val Leu Gly Leu Tyr Leu Ile Arg Phe Ala Glu Val Val Glu Glu
98          245          250          255
100 Ala Cys Thr Asn Leu Leu Pro Asn Val Val Cys Glu Tyr Leu Tyr Asn
101          260          265          270
103 Leu Ser Glu Met Phe Thr Lys Phe Tyr Thr Asn Cys Gln Val Val Gly
104          275          280          285
106 Ser Pro Glu Glu Thr Ser Arg Leu Leu Leu Cys Gln Ala Thr Ala Val
107          290          295          300
109 Val Met Arg Gln Cys Phe Asn Leu Leu Gly Ile Thr Pro Val Tyr Lys
110 305          310          315          320
112 Leu
115 <210> SEQ ID NO: 3
116 <211> LENGTH: 2019
117 <212> TYPE: DNA
118 <213> ORGANISM: Oryza sativa
120 <400> SEQUENCE: 3
121 gcacgagctt acagctgaac acagtgtaca gagcgtcgag caacaactct gtacattaat . 60
122 cacatcttcc ctgctgcaa cagtacctga cctggatgtg gaaccgatgc ttgaagtctc 120
123 aaaaccaggt tttggggatt accagtgcaa caatgctatg agtgtatttt caagaataag 180
124 aggatccgca acaaacttcc gtaaccccat ggcagttggg caggcaattg caaataacct 240
125 cccccagtca aatattatcg aatccatctc tgttgctgga cctggttaca ttaacataac 300
126 gttatccagc aattggattg cacagaggat acaagacatg cttgtttgtg gtatcaaaac 360
127 atgggcacca atcttacctg ttaagagggc agtgctggat tttcatccc ctaatatgtc 420

```

RAW SEQUENCE LISTING

DATE: 04/03/2003

PATENT APPLICATION: US/09/831,683B

TIME: 14:25:52

Input Set : A:\BB1270 USPCT Seq List Rev Feb 03.txt

Output Set: N:\CRF4\04032003\I831683B.raw

```

128 aaaagagatg catgttggac atataaggtc caccataatt ggagatactc tagctcacat 480
129 gtttgagttc acaaatgttg aagttcttcg acgtaacccat gtgggagact ggggtacaca 540
130 gtttgggatg ttgatagagt ttctgtttga gcaattccca gattgggagg atgttgggaa 600
131 ccaggctggt ggagatcttc agagcttcta caaggcatcc aagaaaagat ttgacgatga 660
132 tcctgatttt aaggagaggg ctcggaagc ggtagttcga ctgcaggagg gagaagataa 720
133 ataccgagct gcttggaata aaatatgtca aatcagccga atggagtttg atttggtata 780
134 caaacgcctt aatgtgaagc ttgaagaaaa gggagagagc ttttacaacc cctacattcc 840
135 acctgttttg gaggaattga ctaacaaagg tttgattgtg gaaagtaaag gcgctcgagt 900
136 aatattcggt gaagaccacc ctttgatagt gattaaacaa gatggcggct tcaactatgc 960
137 ctccacagac ttggcagctc tttggtatcg gcttaatgtg gagaaggcag aatggataat 1020
138 atatgtaacg gatgtaggtc agcaacgaca ctttcatatg ttgttctact ctgcaaagat 1080
139 ggctggctgg ctcccagaac aaaatggaaa gaaatacccg aaagcaagcc atgttggatt 1140
140 tggcctagtt cttggttcag atggcaagcg cttccggact cgttggttctg aagttgttcg 1200
141 actggttgat ctacttgatg aggctaaagc tcggagcaaa gcacaactca tcaaacgttt 1260
142 cactggaaat ggtcaaattg ctgactggac agatgatgag ctcgatagga cttcagaggc 1320
143 tataggatat ggtgctgtta agtattcaga tcttaaaaaac aatcggtcga cagactacac 1380
144 atttagtttt gatcaaatgc tgagtgcaca gggaaatact gctgtctacc ttcagtatgc 1440
145 acatgcccggt atctgttcca ttatcaggaa agccagcaag gatgtagaga agttaaaaaat 1500
146 gactggagcc attacccttg gccatccata cgagcgtttc ctcggattac atctcatcca 1560
147 gtttaccgag gttgtggagc aggcttgtgc cgatttacag ccccatcggt tgtgcgacta 1620
148 cttgtatagc ttatccttaa cattctccaa gttttacaca aactgccagg tggttggttc 1680
149 acctgaagaa acgagccgct tgctgctatg tgaagcaaca ggcatcatca tgaggcagtg 1740
150 tttccacctg ttgggcataa caccagtgcg caagctatga caatccacgc cccaatacaa 1800
151 tgccatttgg aagaatttcc aagctataaa tgtaaatagt atattacctt aaaagctaatt 1860
152 gtaaatattg agtgggtgta gtgtcttgta aataggcggg ggctgtaagg cctcgccatc 1920
153 tctgtacatt cttcaatttt ttaataact acggctcgcg ttctttgccg tccctacgaa 1980
154 aataaaaaaa aataatgtaa aaaaaaaaaa aaaaaaaaaa 2019

```

156 <210> SEQ ID NO: 4

157 <211> LENGTH: 587

158 <212> TYPE: PRT

159 <213> ORGANISM: Oryza sativa

161 <400> SEQUENCE: 4

```

162 Glu His Ser Val Gln Ser Val Glu Gln Gln Leu Cys Thr Leu Ile Thr
163 1 5 10 15
165 Ser Ser Leu Arg Ala Thr Val Pro Asp Leu Asp Val Glu Pro Met Leu
166 20 25 30
168 Glu Val Ser Lys Pro Gly Phe Gly Asp Tyr Gln Cys Asn Asn Ala Met
169 35 40 45
171 Ser Val Phe Ser Arg Ile Arg Gly Ser Ala Thr Asn Phe Arg Asn Pro
172 50 55 60
174 Met Ala Val Gly Gln Ala Ile Ala Asn Asn Leu Pro Gln Ser Asn Ile
175 65 70 75 80
177 Ile Glu Ser Ile Ser Val Ala Gly Pro Gly Tyr Ile Asn Ile Thr Leu
178 85 90 95
180 Ser Ser Asn Trp Ile Ala Gln Arg Ile Gln Asp Met Leu Val Cys Gly
181 100 105 110
183 Ile Lys Thr Trp Ala Pro Ile Leu Pro Val Lys Arg Ala Val Leu Asp
184 115 120 125
186 Phe Ser Ser Pro Asn Ile Ala Lys Glu Met His Val Gly His Ile Arg

```

RAW SEQUENCE LISTING

DATE: 04/03/2003

PATENT APPLICATION: US/09/831,683B

TIME: 14:25:52

Input Set : A:\BB1270 USPCT Seq List Rev Feb 03.txt

Output Set: N:\CRF4\04032003\I831683B.raw

```

187      130      135      140
189 Ser Thr Ile Ile Gly Asp Thr Leu Ala His Met Phe Glu Phe Thr Asn
190 145      150      155      160
192 Val Glu Val Leu Arg Arg Asn His Val Gly Asp Trp Gly Thr Gln Phe
193      165      170      175
195 Gly Met Leu Ile Glu Phe Leu Phe Glu Gln Phe Pro Asp Trp Glu Asp
196      180      185      190
198 Val Gly Asn Gln Ala Val Gly Asp Leu Gln Ser Phe Tyr Lys Ala Ser
199      195      200      205
201 Lys Lys Arg Phe Asp Asp Asp Pro Asp Phe Lys Glu Arg Ala Arg Gln
202      210      215      220
204 Ala Val Val Arg Leu Gln Gly Gly Glu Asp Lys Tyr Arg Ala Ala Trp
205 225      230      235      240
207 Lys Lys Ile Cys Gln Ile Ser Arg Met Glu Phe Asp Leu Val Tyr Lys
208      245      250      255
210 Arg Leu Asn Val Lys Leu Glu Glu Lys Gly Glu Ser Phe Tyr Asn Pro
211      260      265      270
213 Tyr Ile Pro Pro Val Leu Glu Glu Leu Thr Asn Lys Gly Leu Ile Val
214      275      280      285
216 Glu Ser Lys Gly Ala Arg Val Ile Phe Val Glu Asp His Pro Leu Ile
217      290      295      300
219 Val Ile Lys Gln Asp Gly Gly Phe Asn Tyr Ala Ser Thr Asp Leu Ala
220 305      310      315      320
222 Ala Leu Trp Tyr Arg Leu Asn Val Glu Lys Ala Glu Trp Ile Ile Tyr
223      325      330      335
225 Val Thr Asp Val Gly Gln Gln Arg His Phe His Met Leu Phe Thr Ala
226      340      345      350
228 Ala Lys Met Ala Gly Trp Leu Pro Glu Gln Asn Gly Lys Lys Tyr Pro
229      355      360      365
231 Lys Ala Ser His Val Gly Phe Gly Leu Val Leu Gly Ser Asp Gly Lys
232      370      375      380
234 Arg Phe Arg Thr Arg Cys Ser Glu Val Val Arg Leu Val Asp Leu Leu
235 385      390      395      400
237 Asp Glu Ala Lys Ala Arg Ser Lys Ala Gln Leu Ile Lys Arg Phe Thr
238      405      410      415
240 Gly Asn Gly Gln Ile Ala Asp Trp Thr Asp Asp Glu Leu Asp Arg Thr
241      420      425      430
243 Ser Glu Ala Ile Gly Tyr Gly Ala Val Lys Tyr Ser Asp Leu Lys Asn
244      435      440      445
246 Asn Arg Leu Thr Asp Tyr Thr Phe Ser Phe Asp Gln Met Leu Ser Asp
247      450      455      460
249 Lys Gly Asn Thr Ala Val Tyr Leu Gln Tyr Ala His Ala Arg Ile Cys
250 465      470      475      480
252 Ser Ile Ile Arg Lys Ala Ser Lys Asp Val Glu Lys Leu Lys Met Thr
253      485      490      495
255 Gly Ala Ile Thr Leu Gly His Pro Tyr Glu Arg Phe Leu Gly Leu His
256      500      505      510
258 Leu Ile Gln Phe Thr Glu Val Val Glu Gln Ala Cys Ala Asp Leu Gln
259      515      520      525

```

RAW SEQUENCE LISTING

DATE: 04/03/2003

PATENT APPLICATION: US/09/831,683B

TIME: 14:25:52

Input Set : A:\BB1270 USPCT Seq List Rev Feb 03.txt

Output Set: N:\CRF4\04032003\I831683B.raw

```

261 Pro His Arg Leu Cys Asp Tyr Leu Tyr Ser Leu Ser Leu Thr Phe Ser
262      530      535      540
264 Lys Phe Tyr Thr Asn Cys Gln Val Val Gly Ser Pro Glu Glu Thr Ser
265 545      550      555      560
267 Arg Leu Leu Leu Cys Glu Ala Thr Gly Ile Ile Met Arg Gln Cys Phe
268      565      570      575
270 His Leu Leu Gly Ile Thr Pro Val His Lys Leu
271      580      585
273 <210> SEQ ID NO: 5
274 <211> LENGTH: 1123
275 <212> TYPE: DNA
276 <213> ORGANISM: Glycine max
278 <400> SEQUENCE: 5
279 gcacgaggtt gaggggtgtag acataccact tattgctgtg aaaagagatg gtggctacaa 60
280 ctattttaca actgatctag catcactttt gtatcgctta aatgaagaaa aacttgaatg 120
281 gattgtatat gttacagata ttgggcagca acagcacttt gatatgctat ttaaggccta 180
282 taggcgtgca ggttggttac caaaggatga gaatgcgtat ccaaaatgta ctcatatagg 240
283 ttttggctctt gttcttgggg aagatggaaa acgatttcgg actcgagca gtgaggttgt 300
284 tcgattagtt gatttacttg atgaagctaa aaggcgtgtg aaaattgcc a ttcttgaacg 360
285 tgatacaact aaagatttgt ctgaggagga gatcgagaaa acatccgagg cagttggtta 420
286 tggggctgtt aagtatgctg atttgaagat caacagatta acaaattaca cttcaactt 480
287 tgatcagatg cttaatgaca aggggaatac tgctgtttat ttgctgtatg cacatgctag 540
288 gatctgttcc attatcagga aatctggtta agacatagaa gaagtaaaga gaaatgggaa 600
289 aatagtgttg gatcatgaag atgaacgtgc attggggctt catttgctac aatttcctga 660
290 ggtttttgag gaggcattgca ccaatttggt gcccaatttc ttgtgtgaat acctttacaa 720
291 tttggcagaa atctttacaa aaaaatttta cgctaattgt caggttgttg ggtcgctga 780
292 ggaaaccagt agactcttgc tatgtgaagc aacggtgact gtgatgagac actgctttta 840
293 tctccttgga attgaacatg tatacaggct atgacctata tataagagat tcatatgcaa 900
294 attcttcac agattttttt gggatataca agtataggaa acttcacaat gaaaattgtt 960
295 caggcaaatt cgaccactcc ttcccctctt ccattttgtt aattttattt gagttgtaac 1020
296 ttgtaagaag taaagtaa attttatagc aatattgttg acaaggcccc atgaaaaatt 1080
297 tatttagatg agttatattt cgtttaaaaa aaaaaaaaaa aaa 1123
299 <210> SEQ ID NO: 6
300 <211> LENGTH: 288
301 <212> TYPE: PRT
302 <213> ORGANISM: Glycine max
304 <400> SEQUENCE: 6
305 Val Glu Gly Val Asp Ile Pro Leu Ile Ala Val Lys Arg Asp Gly Gly
306 1 5 10 15
308 Tyr Asn Tyr Phe Thr Thr Asp Leu Ala Ser Leu Trp Tyr Arg Leu Asn
309 20 25 30
311 Glu Glu Lys Leu Glu Trp Ile Val Tyr Val Thr Asp Ile Gly Gln Gln
312 35 40 45
314 Gln His Phe Asp Met Leu Phe Lys Ala Tyr Arg Arg Ala Gly Trp Leu
315 50 55 60
317 Pro Lys Asp Glu Asn Ala Tyr Pro Lys Cys Thr His Ile Gly Phe Gly
318 65 70 75 80
320 Leu Val Leu Gly Glu Asp Gly Lys Arg Phe Arg Thr Arg Ser Ser Glu
321 85 90 95

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/831,683B

DATE: 04/03/2003
TIME: 14:25:53

Input Set : A:\BB1270 USPCT Seq List Rev Feb 03.txt
Output Set: N:\CRF4\04032003\I831683B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:11; N Pos. 139,238,431
Seq#:12; Xaa Pos. 130
Seq#:15; N Pos. 14,250,293,341,350,351,383,399,401
Seq#:16; Xaa Pos. 55,69
Seq#:19; N Pos. 18,610,713,720,740
Seq#:20; Xaa Pos. 6,203,235,238
Seq#:23; N Pos. 486,677,742,810,824,851,889,893,910
Seq#:24; Xaa Pos. 96,156,220
Seq#:25; N Pos. 274,391,398,407,429,449,454,466,471,475,488,494,497,506,513
Seq#:25; N Pos. 517,520,530,541,546
Seq#:26; Xaa Pos. 51
Seq#:27; N Pos. 18,35,51,159,165,281,286,288,325,342,373
Seq#:28; Xaa Pos. 6,12,17,53,55,94,96,109
Seq#:29; N Pos. 350,378,408,414,432,433,444,452,458,466,487,494,503,523,535
Seq#:29; N Pos. 539,548,555
Seq#:31; N Pos. 448,512
Seq#:33; N Pos. 386,423,459,481,483,486
Seq#:34; Xaa Pos. 63
Seq#:35; N Pos. 18,483,505,506
Seq#:37; N Pos. 140,370,411,413,469,481,504,515,530,551,556,564
Seq#:38; Xaa Pos. 38